

The Importance of Past MJO Activity in Determining the Future State of the Midlatitude Circulation

KAI-CHIH TSENG, ELIZABETH A. BARNES, AND ERIC MALONEY

Colorado State University, Fort Collins, Colorado

(Manuscript received 12 July 2019, in final form 6 November 2019)

ABSTRACT

The Madden–Julian oscillation (MJO) is one of the most important sources of predictability on subseasonal to seasonal (S2S) time scales. Many previous studies have explored the impact of the present state of the MJO on the future evolution and predictability of extratropical weather patterns. What is still unclear, however, is the importance of the accumulated influence of past MJO activity on these results. In this study, the importance of past MJO activity in determining the future state of extratropical circulations is examined by using a linear baroclinic model (LBM) and one of the simplest machine learning algorithms: logistic regression. By increasing the complexity of the logistic regression model with additional information about the past activity of the MJO, it is demonstrated that the past 15 days play a dominant role in determining the state of MJO teleconnections more than 15 days into the future. This conclusion is supported by numerical LBM simulations. It is further shown that the past 15 days of additional information are only important for some MJO phases/lead times and not others, and the physical basis for this result is explored.

1. Introduction and motivation

The Madden–Julian oscillation (MJO) is one of the dominant forms of precipitation and circulation variability in the tropics, and propagates eastward along the equator with a life cycle of 20–100 days (Adames and Kim 2016; Madden and Julian 1971). The MJO's influence is felt across the world. In the tropics, the MJO can influence the initiation of El Niño–Southern Oscillation (ENSO) events through the occurrence of westerly wind bursts (Moore and Kleeman 1999), modulate tropical cyclone genesis (Ching et al. 2010; Maloney and Hartmann 2000), and influence precipitation associated with the Asian and Australian monsoon (Lawrence and Webster 2002; Wheeler et al. 2009). In the extratropics, the MJO has been shown to modulate atmospheric river activity along the west coast of the United States (Mundhenk et al. 2016), the frequency of tornadoes and hailstorms in the U.S. Great Plains (Baggett et al. 2018), extreme cold air outbreaks across North America (Lin 2018), and precipitation over New Zealand (Fauchereau et al. 2016) and Brazil (De Souza and Ambrizzi 2006) via its tropical–extratropical teleconnections. Due to the MJO's ability to impact global circulations, and its

intraseasonal time scales, the MJO has been regarded as one of the most important sources of predictability on subseasonal to seasonal (S2S) time scales (2–5 weeks; Hamill and Kiladis 2014).

In the past decade, multiple studies have leveraged the MJO for S2S prediction by developing empirical models based on the geographical location and the amplitude of MJO forcing (i.e., MJO indices; Cassou 2008; Baggett et al. 2018; Mundhenk et al. 2018). Most of these empirical models are based on MJO information at the time of the forecast (i.e., the MJO at lag 0). However, the MJO's influence on the atmosphere takes time to develop, especially in the extratropical regions far from the heat source (Mori and Watanabe 2008). By analyzing observational data and employing numerical experiments, Mori and Watanabe (2008) and Tseng et al. (2019) demonstrated the extratropical Pacific–North America (PNA) pattern generated by MJO heating takes around 10–15 days to develop, indicating that prediction of the teleconnection patterns at early forecast lead times may rely on MJO information before lag 0. We define the lag (lead) as the number of days before (after) the present state; for example, lag -1 indicates 1 day before the current state (day 0). In addition, the teleconnections generated by past MJO forcing can interfere with the teleconnections generated

Corresponding author: Kai-Chih Tseng, kctsens@rams.colostate.edu

DOI: 10.1175/JCLI-D-19-0512.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

by later MJO forcing due to the memory of atmosphere. For example, previous studies have shown that particular MJO phases (i.e., phases 2, 3, 6, and 7) favor a more robust PNA-like pattern while the teleconnection patterns generated by the other phases are less consistent from one event to the next (Tseng et al. 2019). Thus, the superimpositions of different teleconnection signals from MJO forcing at different time lags likely jointly help determine the extratropical circulation. These two factors imply that developing empirical models based solely on the present state of the MJO (i.e., MJO at lag 0) may be improved by incorporating information about the past MJO states as well. The importance of the accumulated impacts of past MJO activity is also relevant for diagnostic studies. For example, Cassou (2008) demonstrated that the MJO can significantly influence the North Atlantic oscillation (NAO) with a lead time of 0–14 days. However, MJO teleconnections take time to develop over the Atlantic far from the heat source, and thus modulation of the NAO at 0–14 days lead is likely also impacted by the past state, in addition to the present state, of the MJO. Although the concept of an empirical prediction may not be directly relevant to those from dynamical models (where information is included in the three-dimensional atmospheric fields), the insights obtained from these simple empirical models can help us better understand initial condition-dependent forecast skill (e.g., conditional forecasts).

In this study, we aim to quantify the influence of past MJO activity on the future state of the extratropical circulation. To do this, we focus on three subquestions: 1) Can an empirical logistic regression model be improved by including information about the MJO before lag 0? 2) At what point in the past does the additional information about the MJO no longer benefit the empirical model? 3) What physical mechanisms explain why additional lags are necessary? By training one of the simplest machine learning algorithms, a logistic regression model, we will demonstrate that MJO activity before lag 0 plays a critical role in determining future midlatitude geopotential height anomalies. The logistic regression model allows for minor nonlinearity, but results for linear regression are shown in appendix B as an additional reference. In addition, we employ numerical experiments in a linear baroclinic model (LBM) and use the dynamics of MJO teleconnection to support what we found in the logistic regression model.

The manuscript is organized as follows. Section 2 provides detailed descriptions of the data, methods, and logistic regression model. In section 3, the results of predicting daily geopotential height anomalies with the logistic regression models are demonstrated using an MJO index as a predictor. We also demonstrate that

different phases require different amounts of MJO information before lag 0 for improving model prediction skills. In section 4, results of the logistic regression model are compared to those from the LBM simulations. In section 5, a physical explanation for the MJO phase dependence of results shown in the end of section 3 is provided. Section 6 is a discussion and conclusions.

2. Data and method

a. Data

Thirty-six years (1979–2015) of daily 500-hPa geopotential height (Z500) data are acquired from the European Centre for Medium-Range Forecasts (ECMWF) interim reanalysis (ERA-Interim; Dee et al. 2011). Since MJO teleconnections are largely characterized by a barotropic structure, the choice of 500 hPa does not qualitatively change the results shown in this study. To derive the anomalous daily Z500, the first three harmonics of the daily climatology and the linear trend are removed. Thus, except for the seasonal cycle, the Z500 anomalies contain the variability spanning daily to interannual time scales. The Z500 field is additionally interpolated onto a $2.5^\circ \times 2.5^\circ$ grid for ease of comparison with the linear baroclinic model results.

The outgoing longwave radiation (OLR) MJO index (OMI) acquired from <https://www.esrl.noaa.gov/psd/mjo/mjoindex> is used as a proxy for MJO convection. The OMI1 and OMI2 indices correspond to the two leading principal components of the equatorial-averaged OLR derived from empirical orthogonal function analysis. To define the MJO phase, the MJO phase angle is defined as $\tan^{-1}[-(\text{OMI1}/\text{OMI2})]$. This angle is then used to approximate the location of MJO convection. A detailed description of the MJO index can be found in Kiladis et al. (2014). Since MJO-induced midlatitude Z500 anomalies are stronger during the boreal winter than the boreal summer, we focus this study on winter days from November to March. An MJO event is defined as when the OMI amplitude is greater than one standard deviation (i.e., $\sqrt{\text{OMI1}^2 + \text{OMI2}^2} > \sigma$).

b. Logistic regression model

We aim to quantify the time scales of influence of past MJO activity on the future state of the extratropical circulation. To do this, we begin by investigating whether an empirical logistic regression model can be improved by including information about the MJO before lag 0. By adding additional prior (before lag 0) MJO information to the input variables, we can examine if the MJO teleconnection prediction skill is improved and quantify the time lags at which these additional predictors no

longer provide benefit. The logistic regression model—one of the most basic forms of an artificial neural networks—has been widely used for classification problems (i.e., “yes/no” predictions; Wilks 2011; Slade and Maloney 2013). The formulation is nearly identical to a linear regression model, except that a nonlinear activation function is applied to the output of the linear regression. The purpose of this activation function is to add nonlinearity to the otherwise linear model.

In this study, we convert the daily Z500 anomaly into a logistic value by mapping each anomaly to its sign. That is, $\text{sign}(Z500) = 1$ when $Z500 \geq 0$ and $\text{sign}(Z500) = 0$ when $Z500 < 0$. Equations (1) and (2) define the logistic regression model which uses OMI1 and OMI2 at lag 0 as predictors to predict the $\text{sign}(Z500)$ τ days into the future.

$$h_{i,\tau} = w_{\text{OMI1},\tau} \times \text{OMI1}_{i,0} + w_{\text{OMI2},\tau} \times \text{OMI2}_{i,0} + b_{\tau}, \quad (1)$$

$$\text{sign}(Z500)_{i,\tau}^{\text{predict}} = \frac{1}{1 + e^{-h_{i,\tau}}}. \quad (2)$$

Equation (1) is the linear part of the logistic model. $\text{OMI1}_{i,0}$ and $\text{OMI2}_{i,0}$ refer to the OMI indices of the i th MJO event at lag 0, $w_{\text{OMI1},\tau}$ and $w_{\text{OMI2},\tau}$ are the corresponding coefficients, b_{τ} is the “bias unit,” which is a constant, and $h_{i,\tau}$ is the predicted geopotential height anomaly by the linear regression model, which has units of meters. In this study, the forecast lead time τ spans from 0 to 30 days. Equation (2) is the sigmoid activation function that converts the output of Eq. (1) (i.e., $h_{i,\tau}$) into logistic values. Thus, plugging $h_{i,\tau}$ into this activation function provides a final prediction of the sign of the Z500 anomaly, $\text{sign}(Z500)_{i,\tau}^{\text{predict}}$. According to the definition of Eq. (2), $\text{sign}(Z500)_{i,\tau}^{\text{predict}}$ is always greater than 0 and smaller than 1. Thus, a decision boundary of 0.5 is used. Specifically, $\text{sign}(Z500)_{i,\tau}^{\text{predict}}$ is considered as 1 when it is greater than/equal to 0.5 and 0 when the value is less than 0.5. For each grid point and forecast lead time, we develop a unique logistic model. In this way, the models over different grid points do not rely on one another. In this study, we use a gradient descent optimizer to determine the optimal values of the coefficients (e.g., $w_{\text{OMI1},\tau}$, $w_{\text{OMI2},\tau}$, and b_{τ}). Detailed information is given in appendix A.

c. Linear baroclinic model

To augment the results from the logistic regression model, we additionally conduct numerical experiments with a linear baroclinic model (Watanabe and Kimoto 2000). Specifically, we compare five simulations, with each of the five simulations forced with the composite heating anomaly from the same MJO events but initialized at different time lags (i.e., lag = 0, -5, -10, -15, -20).

This experimental setup allows us to examine how past values of MJO forcing (before lag 0) impact the generation of particular extratropical teleconnection patterns. These results are then compared to the results based on the logistic regression model. Additional details are provided in section 4.

1) MODEL DESCRIPTION

In the LBM, the primitive equations are linearized about a given basic state and the anomalous response of the circulation is calculated based on the prescribed forcing. The model formulation can be found in the appendix of Watanabe and Kimoto (2000) and the LBM users’ guide at <https://ccsr.aori.u-tokyo.ac.jp/~lbm/lbm/doc2.2.pdf>. Instead of looking for a steady-state extratropical response to the MJO forcing, we utilize the time-integration model configuration where the forcing and response are time-dependent. This configuration is further used to develop the storm track model in Watanabe and Kimoto (2000). Both “time-integration” and “storm track” models share the same formulation but different parameter settings, including forcing and numerical damping. Different from the random white noise forcing used in the storm track model, an eastward propagating MJO heating is used as an forcing to perturb the model in this study (see section 4). Similar to Tseng et al. (2019), the model is run at T42 horizontal resolution ($\sim 2.8^\circ \times 2.8^\circ$) with 20 sigma levels and utilizes linear damping (including Rayleigh friction and Newtonian damping). The e -folding time scale for the numerical damping is 20 days for all vertical layers except the top 1 and bottom 3 layers, which have an e -folding time scale of 0.5 days. We also employ fourth-order biharmonic damping (i.e. ∇^4) with an e -folding time scale of 2 h for the shortest wavelength. The strong damping in the lower troposphere and scale-dependent damping avoid issues with baroclinic instability and stabilize the integration. Henderson et al. (2017) and Tseng et al. (2019) demonstrated that the LBM can simulate MJO teleconnection patterns comparable to the observational reanalysis with this parameter setup [e.g., Fig. 8 in Henderson et al. (2017) and Figs. 4 and 5 in Tseng et al. (2019)]. All of these parameters are fixed throughout this study, ensuring that any differences between simulations is caused by differences in the prescribed forcing.

2) MODEL BASIC STATES AND FORCING

The LBM requires two inputs: 1) the basic state, which is associated with the linear operators in the model, and 2) the anomalous forcing, which is used to drive the anomalous circulation in the model. For the basic state, three variables acquired from ERA-Interim are used:

surface pressure, horizontal momentum (both zonal and meridional), and temperature. These variables are averaged over the boreal winter (November–March) from 1979–2015 and remain fixed throughout this study.

For the anomalous heating, the daily apparent heat source is calculated based on Yanai et al. (1973). The apparent heat source is defined by the budget residual of the thermodynamic energy equation:

$$Q_1 \equiv \frac{Ds}{Dt} \cong Q_R + Q_c - \frac{\partial \overline{s' \omega'}}{\partial p}. \quad (3)$$

In Eq. (3), s is the dry static energy, which is defined as $c_p T + gz$ where c_p is the specific heat, T is the temperature, g is the gravitational acceleration, and z is the geopotential height. Also, Q_R and Q_c are the energy flux convergence of radiation (both longwave and shortwave) and latent heat. $-(\partial \overline{s' \omega'} / \partial p)$ is the flux convergence of dry static energy by subgrid-scale processes (e.g., cumulus convection and boundary layer turbulence). In this study, Q_1 anomalies are derived by removing the linear warming trend and the daily climatology of the first three harmonics of annual cycle. A Lanczos band-pass filter (20–100 days) is then applied to the anomalous data to get the intraseasonal Q_1 anomaly.

3. Prediction of midlatitude Z500 with logistic regression

In this section, we ask whether an empirical logistic regression model for predicting midlatitude Z500 anomalies can be improved by including information about the MJO before lag 0. We then quantify the particular time lag in the past beyond which additional information about the MJO no longer benefits the model. To test how the MJO teleconnection prediction skill changes with increased past MJO index information in the logistic regression model, we rewrite Eq. (1) into a more general form:

$$h_{i,\tau} = \sum_k [w_{\text{OMI}1_{k,\tau}} \times \text{OMI}1_{i,k} + w_{\text{OMI}2_{k,\tau}} \times \text{OMI}2_{i,k}] + b, \quad (4)$$

$$k \in \mathbb{Z}_0^-,$$

where k is any time lag before lag 0 (i.e., nonpositive integers) and τ is the forecast lead time spanning from 0 to 30 days. For example, values of $k = -10$ and $\tau = 10$ indicate that we are using MJO information from lag 0 to lag -10 to predict the teleconnection patterns 10 days after day 0.

Figure 1 demonstrates how often the logistic regression model with $k = 0$ (i.e., no past MJO information as a predictor) successfully predicts the sign of daily Z500 at various lead times. Dotted regions indicate frequencies that are significantly higher than that from a random

forecast (i.e., 50% chance of forecasting positive or negative sign) at the 95% significance level based on a binomial test. The events are divided into groups according to MJO phase (rows) and the forecast lead time (columns). Darker colors indicate higher success rates. The regions of high success rate (i.e., phase 2, with 10–14-day lead time, or phase 3, with 5–9-day lead time) represent a PNA-like wave train extending from the North Pacific to the Gulf of Alaska. Following phase 3, the PNA-like signal initiates from the extratropical Pacific and then strengthens over the Gulf of Alaska, consistent with the propagation of the stationary Rossby wave generated by MJO convection (Tseng et al. 2018). Since the MJO is a circumnavigating system along the equator, with two adjacent phases typically separated by approximately 5 days, we might expect the predicted Z500 in phase 2 at 10–14-day lead time (Fig. 2f) to be similar to the predicted Z500 in phase 3 at 5–9-day lead time (Fig. 2h). It is because of this that we observe similar patterns along the diagonal directions of Fig. 1 (e.g., Figs. 1j, 1h, and 1f).

To investigate how increasing the MJO information before lag 0 influences the MJO teleconnection prediction skill, Fig. 2 (top panel) shows the loss from the logistic regression models over multiple values of k (i.e., $k = 0, -5, -10, -15, -20$) at one particular grid point ($60^\circ\text{N}, 150^\circ\text{E}$; “×” in Fig. 1). The loss function is defined in appendix A, which can be considered as an analog of the root-mean-square error used in linear regression. Blue shading indicates small loss, or better prediction skill, while red shading indicates high loss, or worse prediction skill. Dotted regions indicate that the cross-entropy is significantly higher than that from a random forecast. Specifically, we randomly sampled MJO events from reanalysis based on the sample size in each MJO phase and lead time. This group of data is considered a “random forecast.” We calculate the cross-entropy based on Eq. (A1) by using this random forecast. The random sampling is then repeated 1000 times to approximate the distribution of cross-entropy values by random forecasts. As long as the value shown in the upper panel of Fig. 2 is significantly higher/lower at the 95% confidence level than the random forecast, the shading is marked with a black dot. Phase/lead times with small loss extend along the diagonal direction of the panels in the upper panel of Fig. 2, and these regions are collocated with the lead times of high success rate in Fig. 1. Looking closely at Fig. 2, one finds that the blue shading darkens as more MJO information is added as a predictor. This is more easily seen in lower panel of Fig. 2, which shows the difference in loss function between the logistic regression model with extra MJO information before lag 0 (i.e., various values of k) and the

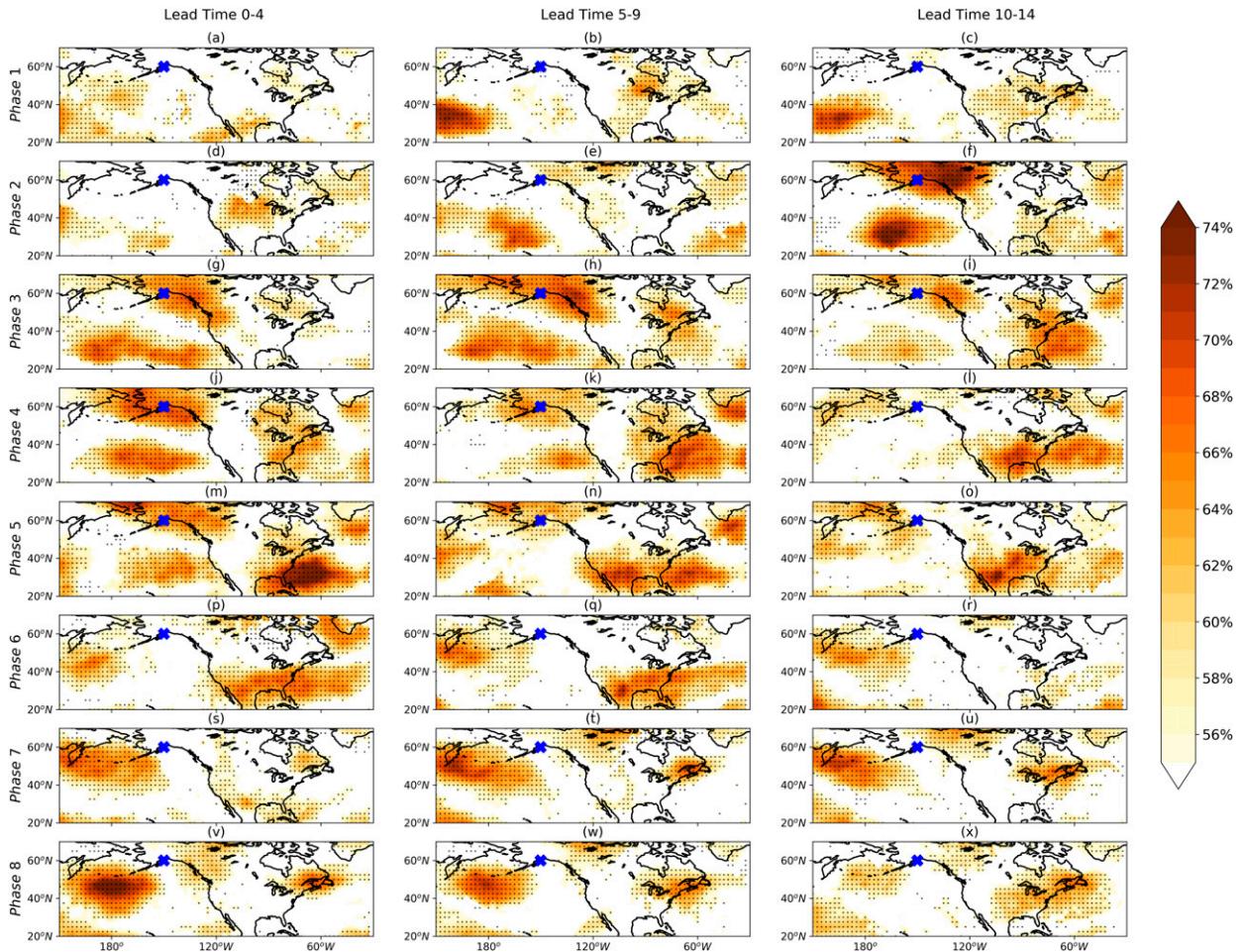


FIG. 1. Frequency (in percent) of correctly predicting the sign of daily Z500 anomalies by the logistic regression model for $k = 0$. A blue “X” denotes the location shown in Figs. 2 and 3. Dotted regions indicate frequencies that are significantly higher than the random forecast at 95% level based on binomial test. The degree of freedom depends on the sample size in each panel.

model only with MJO information from lag 0 (i.e., $k = 0$). The blue color in the lower panel of Fig. 2 indicates the cross-entropy is reduced (or improved prediction skill) with the additional MJO information. Dotted regions indicate the difference in cross-entropy between the two given models is significantly different from zero at the 95% confidence level based on a t test. MJO information from lags 0–5 does not benefit the model prediction skill very much, although the model loss is reduced by 10%–15% when information from lags 0–15 are used (e.g., Fig. 2h).

One possible explanation for why lags 0–5 days offer little additional skill is that these days provide similar information to lag 0 since the MJO is a slow-varying system. In the cases where we provide information before lag –5 days (e.g., Figs. 2g,h,i), MJO activity at these time lags can differ significantly from that at lag 0. In this case, the logistic regression model has additional and nonredundant information that increases prediction

skill. Of course, this increase in skill does not continue without limit, and comparing Figs. 2h and 2i, MJO teleconnection prediction skill barely improves when extra information from lag –15 to lag –20 is included as predictors. This implies that giving the model MJO information from lag 0 to lag –15 may maximize the improvement of prediction skill, while the additional information before lag –15 has little impact on prediction. We will revisit this result with the LBM simulations in section 4. Finally, the lower panel of Fig. 2 presents a perhaps surprising result that not every MJO phase/lead time is improved when additional past MJO information is added to the logistic regression model (e.g., MJO phase 5, 10–15 days forecast lead time in Figs. 2g–i). The mechanism responsible for this feature will be addressed in section 5.

To verify that the results shown in Fig. 2 are not unique to one grid point, Fig. 3 shows maps of the difference in loss between the model for $k = -20$ and the

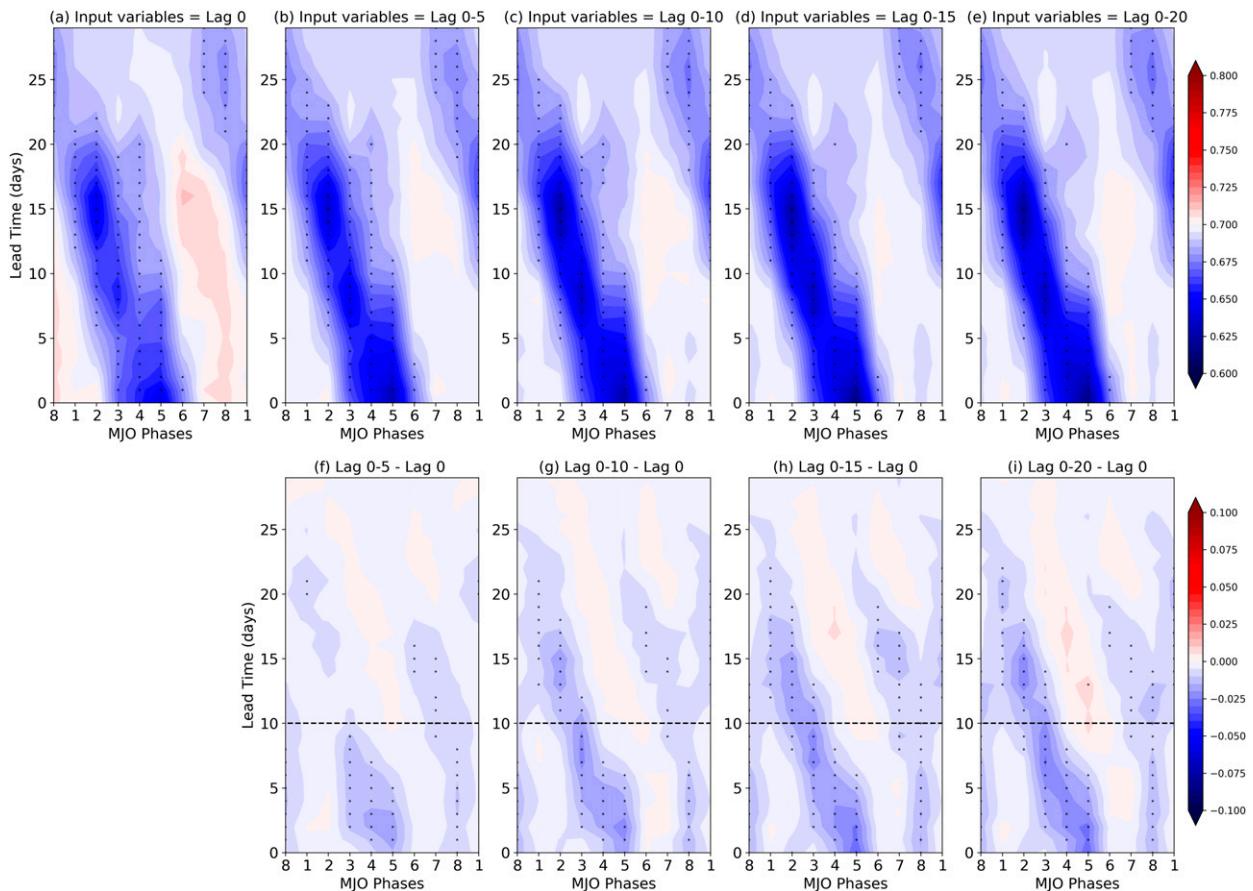


FIG. 2. (top) The average loss (cross-entropy) of the logistic regression model for (a) $k = 0$, (b) $k = -5$, (c) $k = -10$, (d) $k = -15$, and (e) $k = -20$ as written in Eq. (4) for the grid point 70°N , 150°E (blue “ \times ” in Fig. 1). (bottom) The difference in average loss between logistic regression models for (f) $k = -5$ and $k = 0$, (g) $k = -10$ and $k = 0$, (h) $k = -15$ and $k = 0$, and (i) $k = -20$ and $k = 0$ for the grid point 70°N , 150°E (blue “ \times ” in Fig. 1). Dotted regions in the top panel indicate that the cross-entropy is significantly higher/lower than the random forecast at the 95% confidence level based on a bootstrapping analysis. Details of the bootstrapping analysis are provided in the main text. Dotted regions in the bottom panel indicate that the difference in cross-entropy between models is significantly different from 0 at the 95% significance level based on a t test.

model for $k = 0$ (i.e., $k = -20$ minus $k = 0$). Blue shading indicates that the model loss is smaller (skill is improved) when the model includes more past MJO information, while red shading indicates that the model loss is increased (skill is degraded). Similar to the lower panel of Fig. 2, a t test is applied to examine the statistical significance of entropy differences between models with $k = -20$ and $k = 0$. Dotted regions indicate the difference is significantly different from 0 at the 95% confidence level. Comparing Figs. 1 and 3, regions characterized by reduced loss with additional past MJO information (blue shading in Fig. 3) are spatially collocated with the regions that originally showed high prediction skill (shading in Fig. 1).

We also examine the difference between the logistic regression models’ loss with $k = -15$ and $k = -20$ ($k = -20$ minus $k = -15$; shown in Fig. 4). As previously

discussed for a single grid point, MJO teleconnection prediction skill does not greatly improve with inclusion of MJO information before lag -15 over most of the domain.

With regard to the possibility of model dependence, we also include the analysis based on a linear regression model in appendix B. Specifically, we only use Eq. (1) and drop Eq. (2) for forecasting Z500. Since we now forecast the magnitude of geopotential height instead of the sign, the root-mean-square error is used to quantify model performance. In general, the linear regression model qualitatively shows similar results, where the additional lags of the MJO benefits the forecast of future Z500. However, similar to the logistic regression model, the benefit of prior MJO information saturates around $k = -15$ and $k = -20$ does not show much improvement compared to $k = -15$.

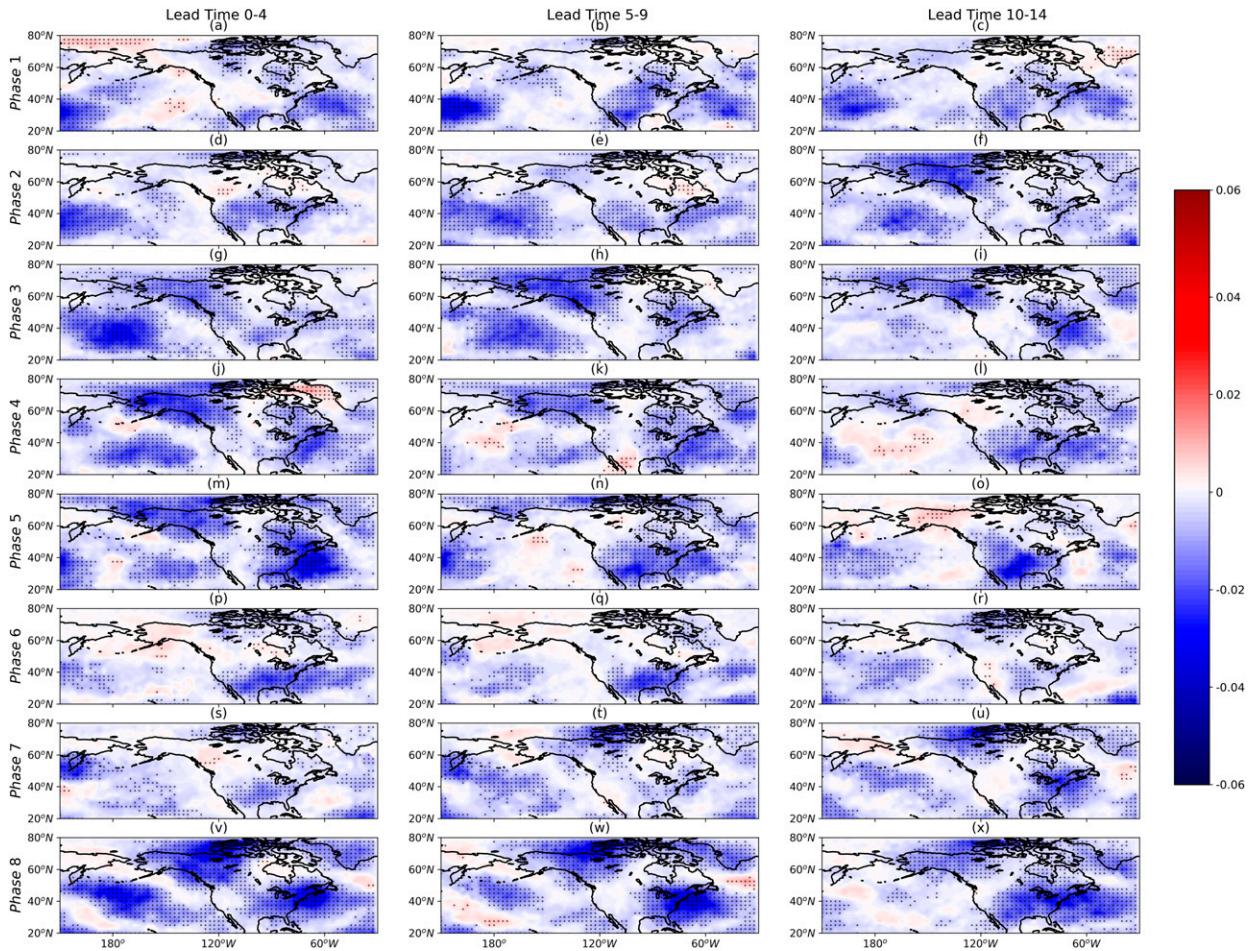


FIG. 3. The difference in average loss (root-mean-square error; unit = m) between logistic regression models for $k = -20$ and $k = 0$. The blue regions indicate reduced cross-entropy with the increase of MJO information (i.e., $k = -20$ minus $k = 0$). Dotted regions indicate that the difference in cross-entropy between $k = -20$ and $k = 0$ is significantly different from 0 at the 95% significance level based on a t test.

4. MJO teleconnections in an LBM

We provide additional physical insight into the results from the logistic regression model using simulations with an LBM. Specifically, we aim to quantify the importance of past MJO information in determining the evolution of the midlatitude geopotential height field.

The MJO forcing for each MJO phase is derived by randomly selecting 20 MJO events from reanalysis data according to the criteria given in section 2 (i.e., the OMI amplitude at lag 0 is greater than 1σ) and calculating the phase-composited Q_1 with respect to lag 0. Five simulations are run for each MJO phase, with each of the five simulations forced with the composite Q_1 from the same MJO events but initialized at different time lags in the past (lag = 0, -5, -10, -15, -20). Figure 5 shows an example of this setup, where the shading represents the equatorial-average (15°S – 15°N) MJO phase 2 forcing for the five simulations. The forcing patterns are identical

from one simulation to the other after lag 0 and the forcing at lag 0 is characterized by the phase 2 heating pattern. What differs between them is the amount of MJO information before lag 0. With this simulation setup, we are able to examine the importance of past MJO activity in driving future midlatitude circulation anomalies. Finally, to increase the robustness of our results, we repeat the five simulations for each phase 24 times, giving us 24 ensemble members per simulation per phase.

The ensemble average Z500 for simulations initialized at lag 0 are shown in Fig. 6 as a function of MJO phase and time after lag 0 (shading). We compare these anomalies with those from the simulation using information up to lag -20 days (in Fig. 5e; contours in Fig. 6). Figure 6 shows large Z500 anomalies concentrated in specific MJO phases rather than along the diagonal directions. As noted previously, the large signals along the diagonal of the MJO phase/lead time plots (such as Figs. 1 and 4) are associated with the propagation of

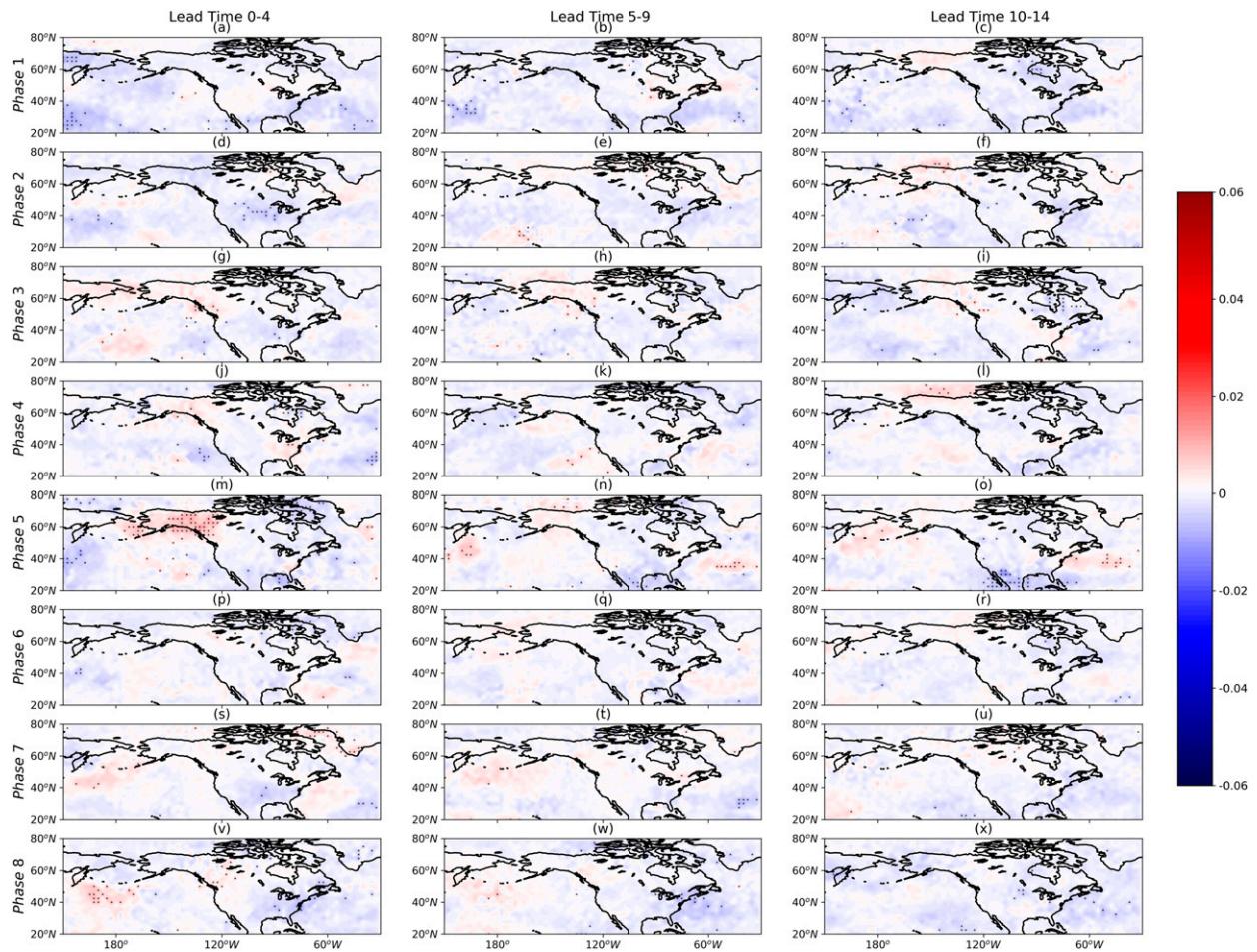


FIG. 4. The difference in average loss (cross-entropy) between logistic regression models for $k = -20$ and $k = -15$. The blue regions indicate reduced cross-entropy with the increase of MJO information (i.e., $k = -20$ minus $k = -15$). Dotted regions indicate that the difference in cross-entropy between $k = -20$ and $k = -15$ is significantly different from 0 at the 95% significance level based on a t test.

MJO convection, where the teleconnections generated by the earlier MJO phases can interfere with the teleconnections generated by later phases. This feature is clear for the simulation that includes the MJO

information from lag 0 to lag -20 (contours in Fig. 6) but is not evident in the simulations without MJO information prior to lag 0 (shading in Fig. 6). Further inspection of Fig. 6 shows that differences between the

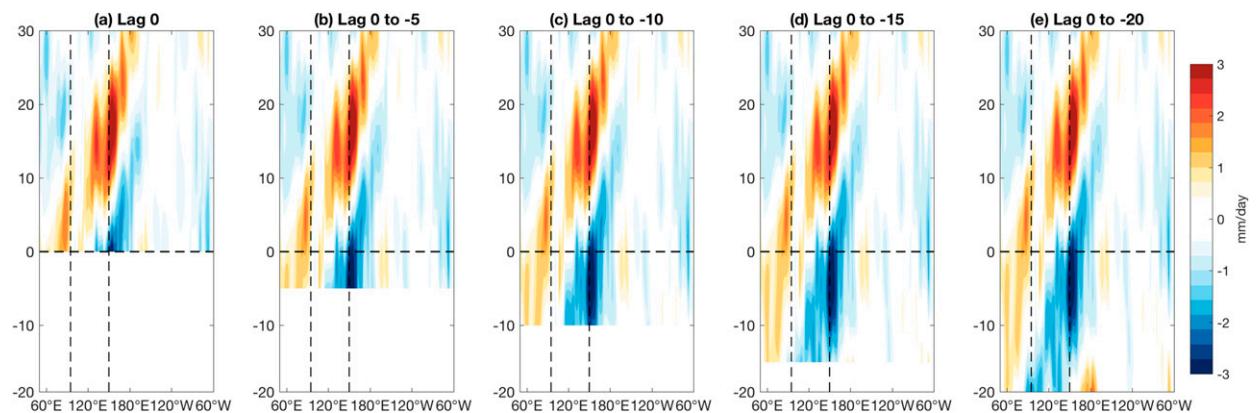


FIG. 5. The equatorial-average (15°S – 15°N) MJO phase 2 column-integrated Q_1 (unit = mm day^{-1}) forcing for initialization of the linear baroclinic model. Each panel denotes the forcing for each of the five different simulations. All of the simulations have identical MJO phase 2 heating at lag 0 and later lags.

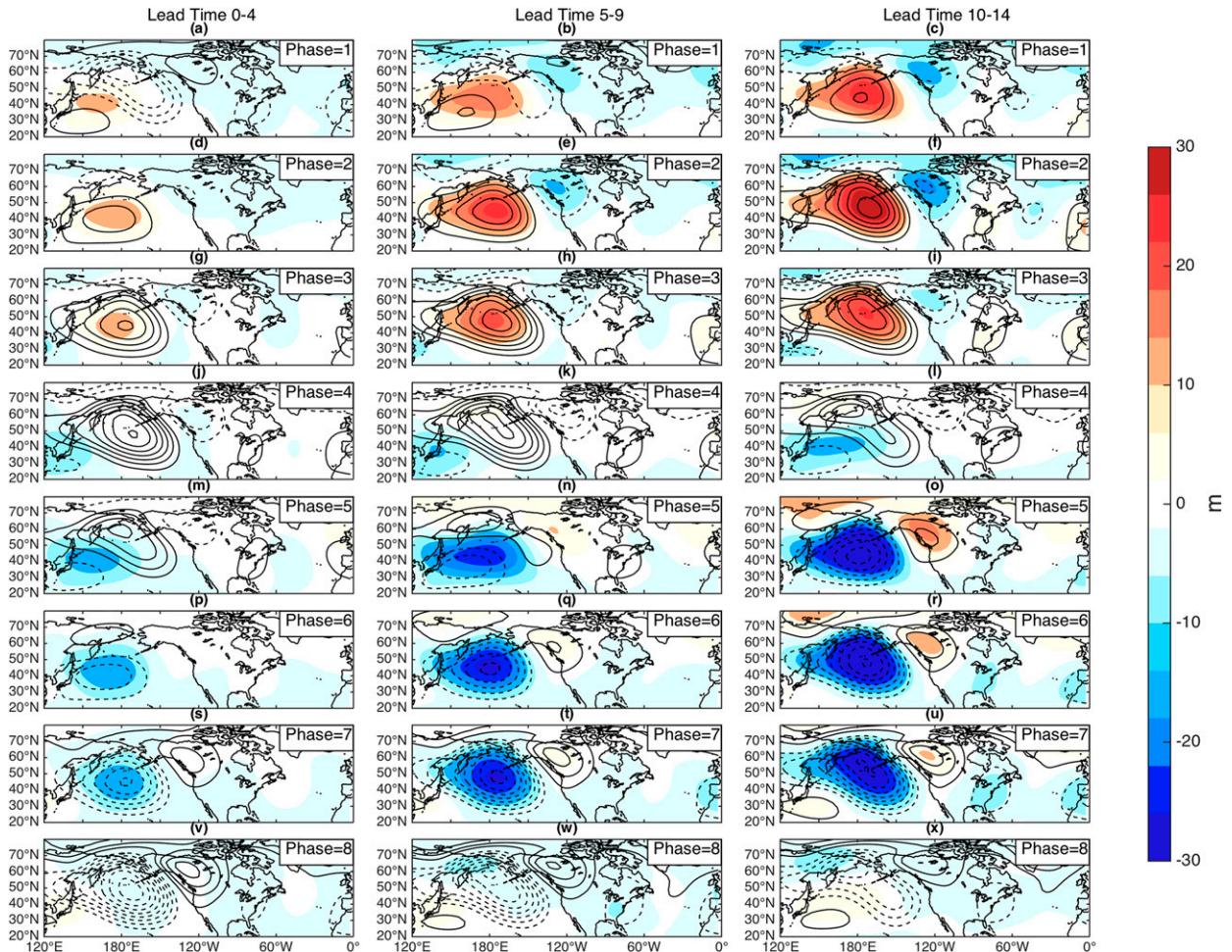


FIG. 6. The ensemble averaged Z500 anomalies for LBM simulations initialized at lag 0 (shading) and initialized at lag -20 (contour). Contours are drawn every 5 m.

shading and the contours are larger for earlier lead times (e.g., first column of Fig. 6, such as Figs. 6g and 6j) and smaller for later lead times (e.g., third column of Fig. 6, such as Figs. 6f and 6i). This can be explained by the time it takes for the MJO teleconnection to develop in the extratropical regions. That is, the MJO teleconnections in the early lead times (e.g., days 0–4) are mostly determined by the MJO forcing before lag 0 while the teleconnections in the later lead times (e.g., days 10–14) are influenced by the MJO forcing after lag 0.

Figure 7 is identical to Fig. 6 except the shading is the ensemble averaged Z500 anomalies from the simulation for lag 0 to -15 (Fig. 5d). The similarity between the shading and contours in Fig. 7 implies that the observed teleconnections are mostly explained by the MJO forcing from lag 0 to lag -15 , while additional information of the MJO before lag -15 has minimal impact on the simulated teleconnection patterns. This result is true

whether one looks at the ensemble average, or each of the 24 ensemble members separately (not shown).

These results support those from the logistic regression model and indicate that 15 days is the approximate time span over which past MJO activity impacts the future evolution of the midlatitude circulation. It is worth mentioning that even though both the LBM and logistic regression models exhibit the same general behavior, the exact meaning of the 15-day time span is slightly different between these two models. We will expand on this point in the next section.

5. Mechanistic explanation

We now address the physical mechanisms that determine the 15-day period over which past MJO activity impacts the future evolution of the midlatitude circulation, with particular emphasis placed on why past

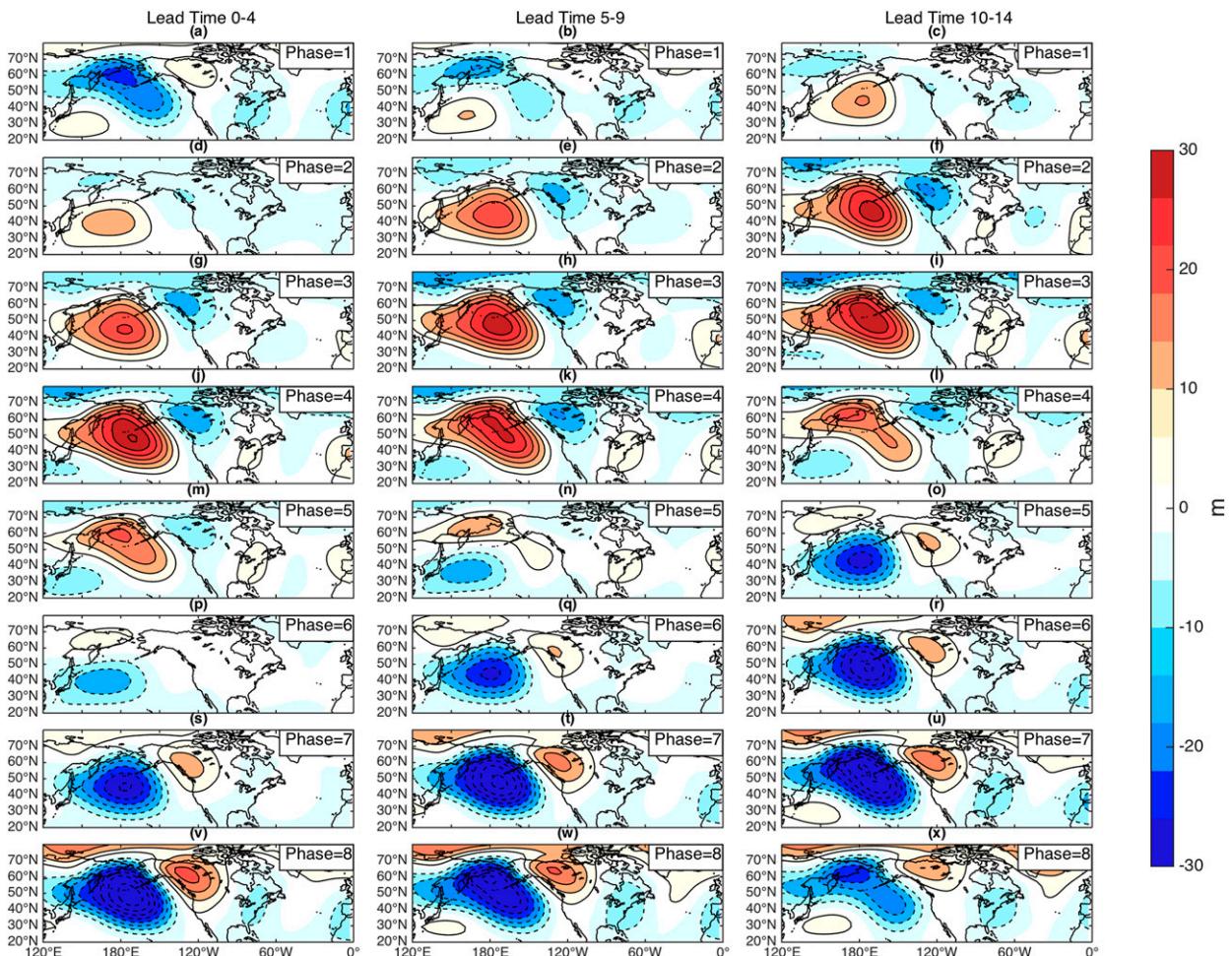


FIG. 7. The ensemble averaged Z500 anomalies for LBM simulations initialized at lag -15 (shading) and initialized at lag -20 (contour). Contours are drawn every 5 m.

information appears to be more important in some phases compared to others (i.e., Fig. 3).

First, we briefly review the literature associated with the dynamics of MJO teleconnections in order to better interpret results from the logistic regression models. According to the numerical tests by Seo and Lee (2017) and Tseng et al. (2019), heating in the eastern Indian Ocean and the western Pacific can generate similar PNA-like teleconnection patterns but with opposite sign. Thus, a PNA-like pattern is more likely to be observed 10–14 days after the MJO phases characterized by dipole heating about the Maritime Continent (e.g., phases 2, 3, 6, and 7) due to constructive interference of the signals excited by the dipole heating in the two regions. On the other hand, the teleconnection patterns in other phases (e.g., phase 1, 4, 5, and 8) tend to vary more from one event to the other because of the destructive interference by the signals in these two regions. Based on these studies, one may hypothesize that a robust

PNA-like signal can be found in the extratropical regions if the MJO heating persists in phases 2 and 3 (or phase 6 and 7), because of the superimposition of the same signals, while a less robust PNA-like pattern can be found if the MJO heating persists in the other phases. We will now explore how this mechanism may influence the “15 additional lags” discussed in the previous sections.

Returning to the results from the logistic regression model, we divide observed MJO events into two groups based on the change in the loss function shown in Fig. 3c. The first group represents the cases where the loss is greatly reduced at 10 days forecast lead time when additional MJO information (i.e., $k = -15$) is included in the model. Specifically, we choose the top one-third of events most improved by the additional past MJO information. The second group represents the cases where the loss shows minimal change at 10 days lead time when the by the additional past MJO information. We choose

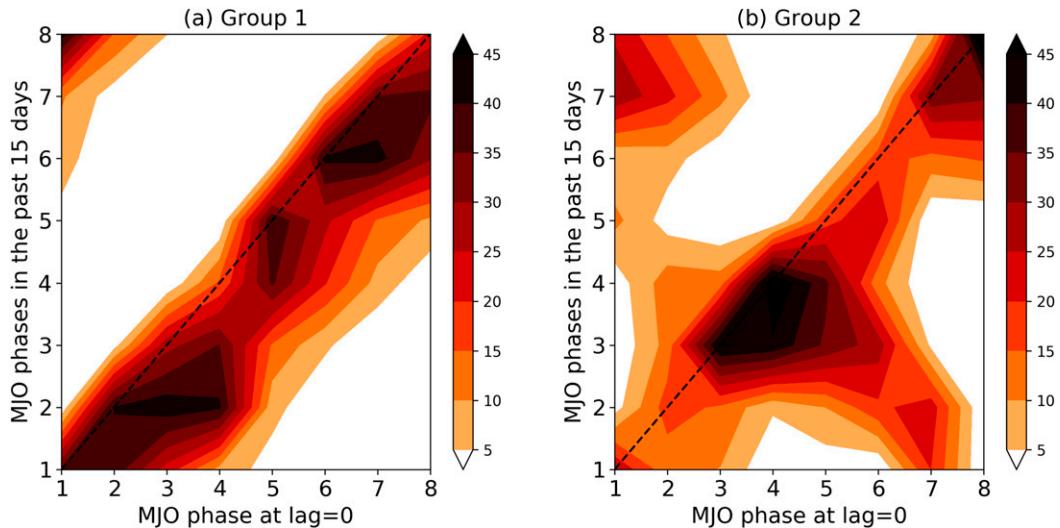


FIG. 8. The frequency of occurrence of the MJO phase for the 15 days before lag 0 (i.e., up to lag -15). The x axis denotes the MJO phase at lag 0, and the y axis denotes the MJO phase over the past 15 days. The MJO events are divided into two different groups as described in the text.

the top one-third events characterized by the least improvement in MJO teleconnection prediction skill. These two groups represent the cases where the concept of 15 additional lags both works (group 1) and fails (group 2).

Figure 8 shows the probability density function of the MJO phase over the 15 days prior to lag 0 for these two groups. The x axis represents the MJO phase at lag 0 and the y axis indicates the MJO phase from lag 0 to lag -15 . Thus, integrating the shading along a constant x axis results in a value of 100%. The diagonal line represents persistent MJO forcing of the same phase over the 15 days. Shading is concentrated below this line for both groups, consistent with eastward propagation of MJO convection (where one phase is followed by the other). For example, if the MJO phase at lag 0 is phase 2, one may expect the MJO phases in the past 15 days to most likely have been phases 1, 8, or 7 rather than phases 3 or 4.

The regions of maximum frequency (shading) in Fig. 8 imply a preference for particular past MJO phases. In group 1, the maximum frequencies appear when the MJO over the past 15 days is in phases 2, 3, 6, and 7. By contrast, the maximum frequencies for group 2 exist when the MJO over the past 15 days is in phases 4 and 8. Based on our previous discussion, phases 2, 3, 6, and 7 are characterized by more consistent teleconnection patterns, while phases 4 and 8 lack consistency. The logistic regression model is only capable of learning consistent information within the training data, and so is only capable of using past phase information when the associated teleconnection patterns are consistent (i.e., group 1). The frequency maxima in Fig. 8a thus signify

the importance of past MJO phases (e.g., phases 2, 3, 6, and 7) in determining the so-called 15 additional lags we found for the logistic regression models. In contrast, the frequency maxima in Fig. 8b signify that some specific MJO phases (e.g., phases 4 and 8) do not provide additional information for prediction. This result is relevant to different types of MJO convective events, including those characterized by fast propagating or standing behavior. In Fig. 8b, there are at least two types of MJO events that can be observed: 1) fast-propagating cases and 2) standing cases. The fast-propagating cases can be found when MJO is phases 2, 3, 6, and 7 at lag 0 (e.g., the cases experience more than four different MJO phases over the past 15 days). On the other hand, the slow-propagating or stationary events can be found when MJO is phases 4, 5, and 8 at lag 0 (e.g., the cases experience fewer than four different phases over the past 15 days). Apparently, both types of MJO can lead to only modest improvement of Z500 predictions. The only feature shared by both types of events is that they both went through the MJO phases generating less-robust teleconnection signals over the past 15 days (e.g., phase 1, 4, 5, and 8). This indicates that the location of tropical heating is the key factor that determines if the additional lags are beneficial for prediction rather than the propagating speed. In general, the conclusion is consistent with Tseng et al. (2019). Tseng et al. (2019) examined similar features in reanalysis based on daily Rossby wave source anomaly in the extratropics, which initially contains no MJO information. Thus, the phase dependence of MJO-induced teleconnections can be regarded as a special case of this mechanism.

Given this discussion, one may be confused about why the LBM does not exhibit a phase dependence in the importance of past information (i.e., no phase is characterized by degraded simulations in LBM when additional lags are used; see Figs. 6 and 7). That is, the LBM results suggest that the MJO activity 15 additional lags in the past is important for the evolution of the midlatitude circulation no matter the present phase of the MJO. The reason for this is that the LBM is a deterministic model, where the future state of extratropical circulations can be considered as the linear superimposition of MJO influence over different lags. Thus, as long as the LBM is forced by Q_1 from the same set of MJO events, the identical response will be simulated regardless of the current MJO phases. However, the logistic regression model based on observations is subject to the predictability of the Earth system, so only consistent relationships can be learned by the logistic regression model, such as the teleconnections generated by MJO phases 2, 3, 6, and 7. This difference explains why the phase/lead time dependence of our results only appears for the logistic regression model rather than the LBM.

6. Conclusions

The impact of past MJO activity on the future state of the extratropical circulation is examined, with particular emphasis placed on quantifying how many days of past information is useful to aid prediction. Using both a logistic regression model and a linear baroclinic model to predict the evolution of midlatitude geopotential height anomalies, we demonstrate that approximately 15 days of prior MJO information importantly impacts the future evolution of the midlatitude circulation. For the logistic regression model, the prediction of MJO teleconnections is improved by including past MJO information, although this improvement is only found for specific phase/lead combinations. We find that only the MJO phases that generate consistent teleconnection patterns (i.e., phases 2, 3, 6, and 7) exhibit the usefulness of these 15 additional lags. This is because the logistic regression model based only on two OMI indices can only learn consistent relationships in the training data.

Our results are based on results from a simple logistic regression model, and it is unclear whether the phase dependence of the results will hold for neural networks

with more complex architectures. In addition, low-frequency climate variability (e.g., ENSO and the quasi-biennial oscillation) is known to modulate the consistency of MJO teleconnections, and thus will likely change the specific phases/lead times for which past MJO information is useful for empirical prediction of the midlatitude circulation. How interannual variability influences how much information on past MJO activity is useful for prediction is an area of research. Regardless, our results suggest that midlatitude empirical prediction schemes based on the MJO may be improved by including information about the past evolution of the MJO.

Acknowledgments. This research has been conducted as part of the NOAA MAPP S2S Prediction Task Force and supported by NOAA Grant NA16OAR4310064 and by the Climate and Large-Scale Dynamics Program of the National Science Foundation under Grant AGS-1841754.

APPENDIX A

Details about the Logistic Regression Model

To evaluate the performance of logistic regression model, the cross-entropy function is used as the loss function in this study [Eq. (A1)].

$$J_\tau = \frac{1}{m} \sum_{i=1}^m \{-y_{i,\tau}^{\text{true}} \times \log[\text{sign}(Z500)_{i,\tau}^{\text{predict}}] - (1 - y_{i,\tau}^{\text{true}}) \times \log[1 - \text{sign}(Z500)_{i,\tau}^{\text{predict}}]\}. \quad (\text{A1})$$

Here J_τ is the “loss” of the logistic regression model, $y_{i,\tau}^{\text{true}}$ denotes the actual (true) value of $\text{sign}(Z500)_{i,\tau}^{\text{predict}}$, and the summation shows that J_τ is derived as an average over m events. Smaller values of J_τ indicate a better model (i.e., a model with lower loss and higher prediction skill). The advantage of using cross-entropy is that the loss (i.e., J_τ) grows exponentially if the model incorrectly predicts the logistic value, providing a larger penalty than the root-mean-square error.

Regularization is a technique used to prevent a trained model from overfitting, which ultimately reduces the error for unseen (out of sample) data. Regularization most often involves adding an additional term to the loss function, as shown below:

$$J_\tau = \frac{1}{m} \sum_{i=1}^m \{-y_{i,\tau}^{\text{true}} \times \log[\text{sign}(Z500)_{i,\tau}^{\text{predict}}] - (1 - y_{i,\tau}^{\text{true}}) \times \log[1 - \text{sign}(Z500)_{i,\tau}^{\text{predict}}]\} + \frac{\lambda}{2m} (w_{\text{OMI}1_0}^2 + w_{\text{OMI}2_0}^2). \quad (\text{A2})$$

Equation (A2) is nearly identical to Eq. (A1) except that it includes an additional term, the regularization term. Note that λ is a positive value (to be explained later in the appendix), and $w_{\text{OMI1},\tau}$ and $w_{\text{OMI2},\tau}$ are the regression coefficients shown in Eq. (1). The regularization term is always positive or zero, with larger values of $w_{\text{OMI1},\tau}$ and $w_{\text{OMI2},\tau}$ leading to a larger loss J_τ . Thus, this regularization term ensures that the coefficients remain small, reducing the chance of overfitting. However, a model with too large a regularization term will force the regression coefficients to near-zero, resulting in a poor predictive model.

We use “gradient descent” to optimize our model to identify the best coefficients for prediction, and the proper amplitude of regularization (i.e., λ). Although gradient descent is the common method for training neural networks, we include its formulation here for completeness. Specifically, the regression coefficients and bias unit are updated during the training process as follows:

$$\begin{aligned}
 w_{\text{OMI1},\tau}^{(j+1)} &= w_{\text{OMI1},\tau}^{(j)} - \eta \frac{\partial J_\tau}{\partial w_{\text{OMI1},\tau}^{(j)}} \\
 w_{\text{OMI2},\tau}^{(j+1)} &= w_{\text{OMI2},\tau}^{(j)} - \eta \frac{\partial J_\tau}{\partial w_{\text{OMI2},\tau}^{(j)}} \\
 b_\tau^{(j+1)} &= b_\tau^{(j)} - \eta \frac{\partial J_\tau}{\partial b_\tau^{(j)}}
 \end{aligned}
 \tag{A3}$$

The superscript $j + 1$ indicates the $(j + 1)$ th step of the training, or updating of the coefficients. The most right-

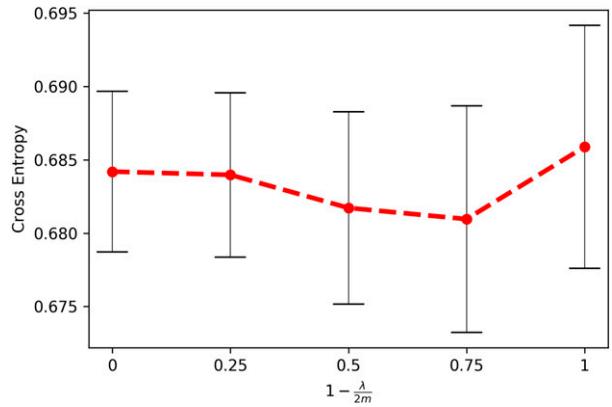


FIG. A1. The average loss (cross-entropy) of the logistic regression model for $k = -20$ averaged over lags 0–30 for 20 different training/testing sets as a function of the regularization term. This example is for the model at the grid point 60°N, 150°E. Error bars denote plus/minus one standard deviation across the 20 training/testing sets.

hand terms are the gradients of the loss function with respect to the coefficients and bias. The minus signs represent the downgradient direction, specifying that the coefficients should be updated in the direction to reduce the loss. Also, η is the learning rate parameter, which indicates the amount to adjust the coefficients with each iteration.

The gradients of the loss function can be analytically derived by plugging Eq. (A2) into Eq. (A3):

$$\begin{aligned}
 w_{\text{OMI1},\tau}^{(j+1)} &= \left(1 - \eta \frac{\lambda}{m}\right) w_{\text{OMI1},\tau}^{(j)} - \eta \frac{1}{m} \sum_{i=1}^m [\text{sign}(Z500)_{i,\tau}^{\text{predict}} - y_{i,\tau}^{\text{true}}] w_{\text{OMI1},\tau}^{(j)} \\
 w_{\text{OMI2},\tau}^{(j+1)} &= \left(1 - \eta \frac{\lambda}{m}\right) w_{\text{OMI2},\tau}^{(j)} - \eta \frac{1}{m} \sum_{i=1}^m [\text{sign}(Z500)_{i,\tau}^{\text{predict}} - y_{i,\tau}^{\text{true}}] w_{\text{OMI2},\tau}^{(j)} \\
 b_\tau^{(j+1)} &= b_\tau^{(j)} - \eta \frac{1}{m} \sum_{i=1}^m [\text{sign}(Z500)_{i,\tau}^{\text{predict}} - y_{i,\tau}^{\text{true}}] b_\tau^{(j)}
 \end{aligned}
 \tag{A4}$$

Using Eq. (A4), one can iteratively train the logistic regression model to optimize the regression coefficients and the bias term. We train over our training set 40 times (40 epochs), and additionally perform 20 cross-validation cycles (i.e., resample our data 20 times) for the training and testing sets. Each training set contains 2/3 of the total data and the rest is used as the testing set. The final loss for a given model is defined as the

average loss for the testing datasets over all 20 cross-validations.

As noted previously, when we increase the number of input variables, the logistic regression model is more likely to get into the condition of overfitting. Thus, a regularization term is used in this study. Here, we use logistic regression models with $k = -20$ and τ spanning from 0 to 30 in Eq. (4) to demonstrate how

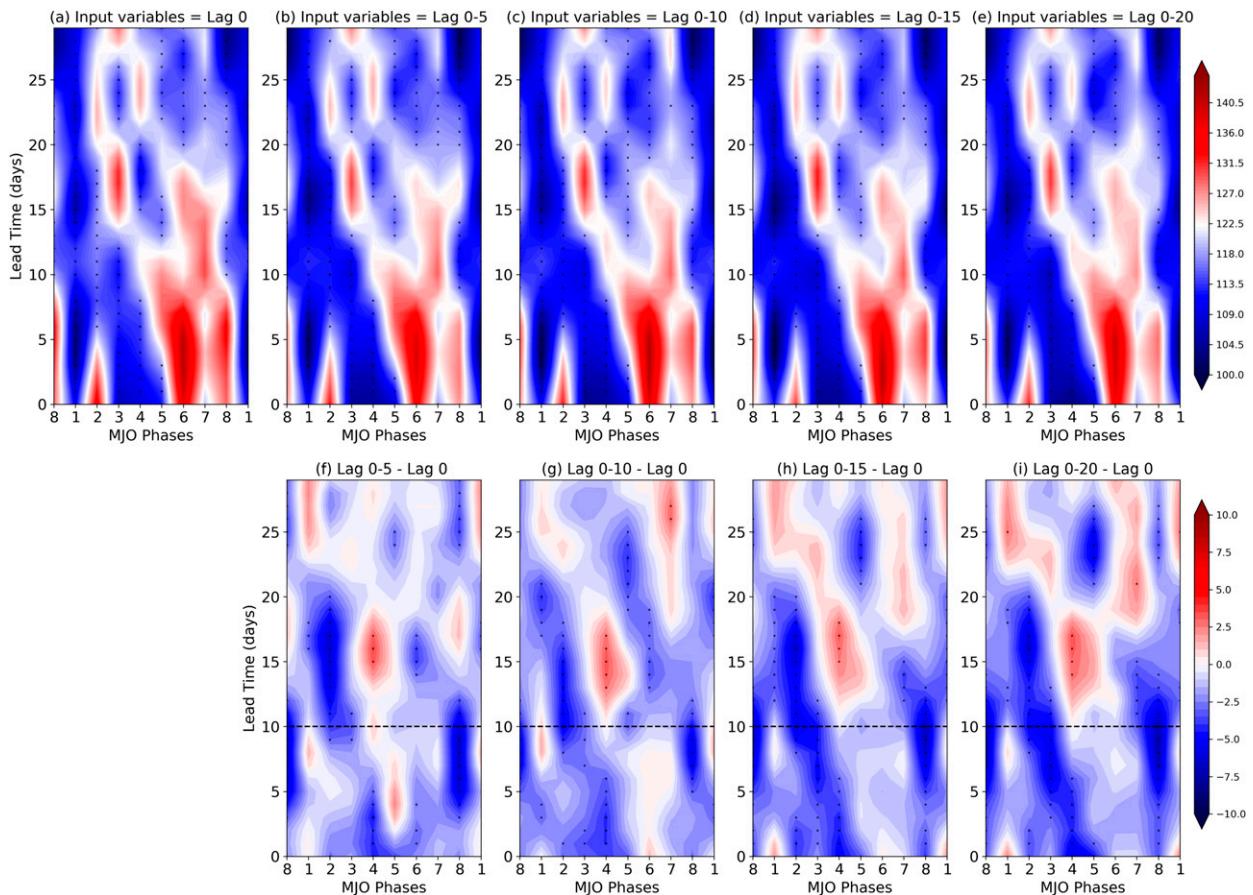


FIG. B1. As in Fig. 2, but using a linear regression model. Shading in the top panel shows the averaged root-mean-square error (unit = m). Dotted regions indicate that the root-mean-square error is significantly higher/lower than the random forecast at the 95% confidence level based on bootstrapping analysis. Shading in the bottom panel shows the difference in root-mean-square error between two given linear regression models (e.g., $k = -20$ minus $k = 0$). The dotted regions indicate the values are significantly different from 0 based on a t test.

the amplitude of regularization influences models' prediction skill.

The choice of the regularization term λ must be chosen before the final training of the model. Figure A1 shows the loss (cross-entropy) of the logistic regression model for $k = -20$ averaged over lags 0–30 for 20 different epochs as a function of the regularization term for one particular grid point. Error bars denote plus/minus one standard deviation of the cross-entropy across the 20 testing sets. Figure A1 shows that the average loss has a local minimum when $[1 - \eta(\lambda/m)] = 0.75$ (or $\eta(\lambda/m) = 0.25$) and higher loss function on both ends. This behavior is consistent with the previous discussion that too large or too small of a regularization term can degrade the performance of the predictive model. Thus, we choose the regularization term that leads to the minimum loss, namely $[1 - \eta(\lambda/m)] = 0.75$. The λ selection in other extratropical

regions and the other k term are based on the same analysis given above.

APPENDIX B

Analysis Based on a Linear Regression Model

In this appendix, we show results utilizing a linear regression model. Specifically, only Eq. (1) is used for forecasting the daily Z500 while Eq. (2) is dropped. In addition, the model forecast skill is defined by the root-mean-square error [Eq. (B1)]:

$$J_{\tau} = \sqrt{\frac{1}{m} \sum_{i=1}^m (h_{i,\tau} - y_{i,\tau}^{\text{true}})^2}. \quad (\text{B1})$$

The notation in Eq. (B1) is identical to that in appendix A. Results from the linear regression model are shown in

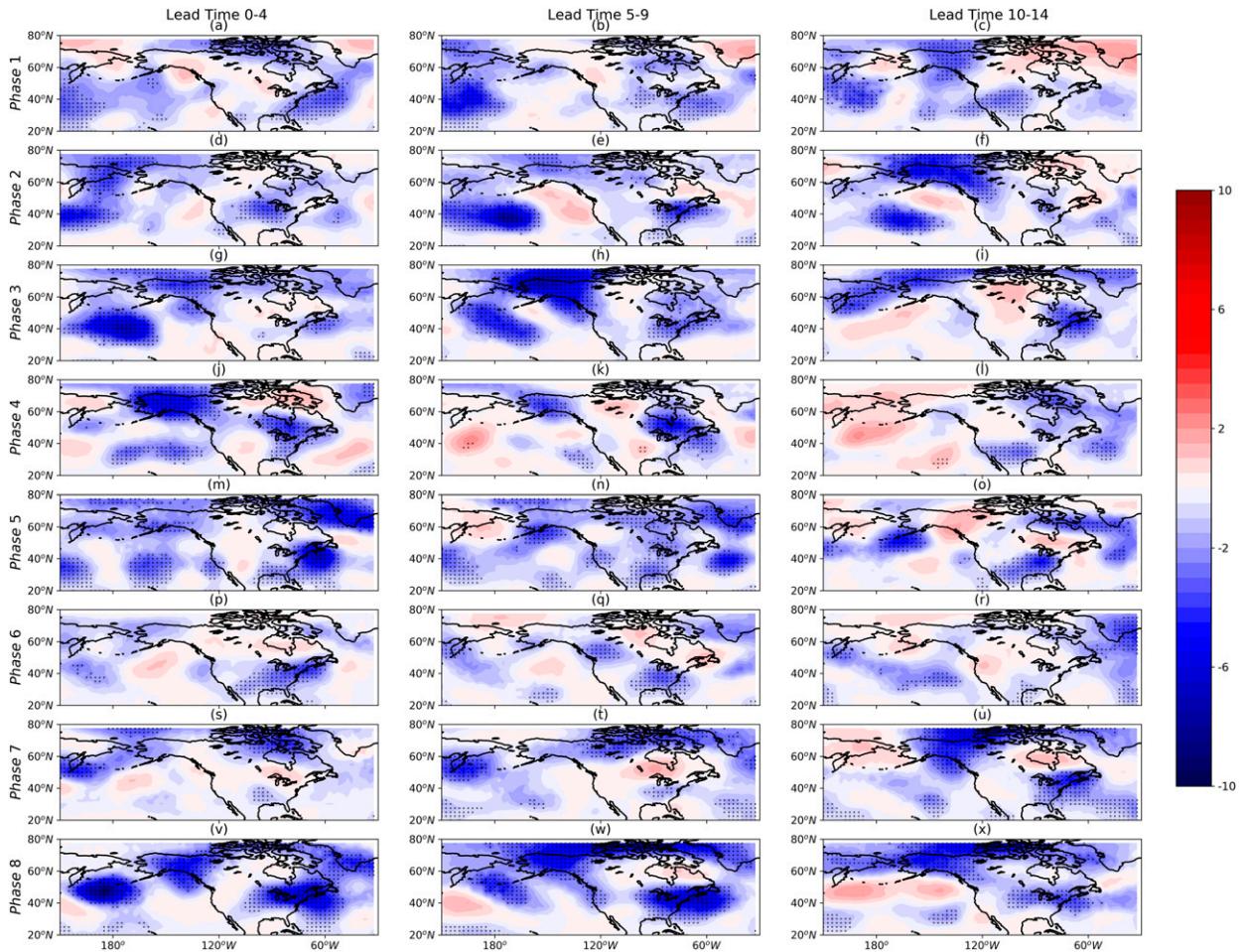


FIG. B2. As in Fig. 3, but using a linear regression model. The blue regions indicate reduced root-mean-square error with the increase of MJO information (i.e., $k = -20$ minus $k = 0$). Dotted regions indicate the difference in root-mean-square error between $k = -20$ and $k = 0$ is significantly different from 0 at the 95% significant level based on a t test.

Figs. B1 and B2. Similar to Fig. 2, the upper panel of Fig. B1 shows the loss (root-mean-square error) from the linear regression models over multiple values of k (i.e., $k = 0, -5, -10, -15, -20$) at one particular grid point ($60^\circ\text{N}, 150^\circ\text{E}$; “ \times ” in Fig. 1). A similar bootstrapping analysis is performed for Fig. B1 to examine the statistical significance, except the distribution of the loss function is based on root-mean-square error. Dotted regions indicate that the cross-entropy is significantly higher/lower than the random forecast at the 95% confidence level. The bottom panel of Fig. B1 shows the difference in root-mean-square error between two given models (e.g., $k = -20$ minus $k = 0$). The dotted regions indicate that the values are significantly different from 0 based on a t test. In general, the results in Fig. B1 are consistent with Fig. 2, where the additional lags of MJO information benefit the prediction of Z500. However,

similar to Fig. 2, the additional skill brought by past MJO information saturates around $k = -15$, while additional information does not show an influence on Z500 prediction skill.

Figure B2 shows maps of differences in the loss function between $k = -20$ and $k = 0$ (i.e., $k = -20$ minus $k = 0$). Similar to Fig. 3, the blue shading indicates the prediction skill is improved (reduced root-mean-square error) when additional lags are used. In Fig. B2, the regions characterized by improved skill are spatially collocated with the regions shown in Fig. 3. We also show the difference in loss function between $k = -20$ and $k = -15$ (i.e., $k = -20$ minus $k = -15$) in Fig. B3. Similar to Fig. 4, most of regions are characterized by insignificant change in the loss function. From Figs. B1 and B3, one can find the linear regression model qualitatively shows the same result as the logistic regression model.

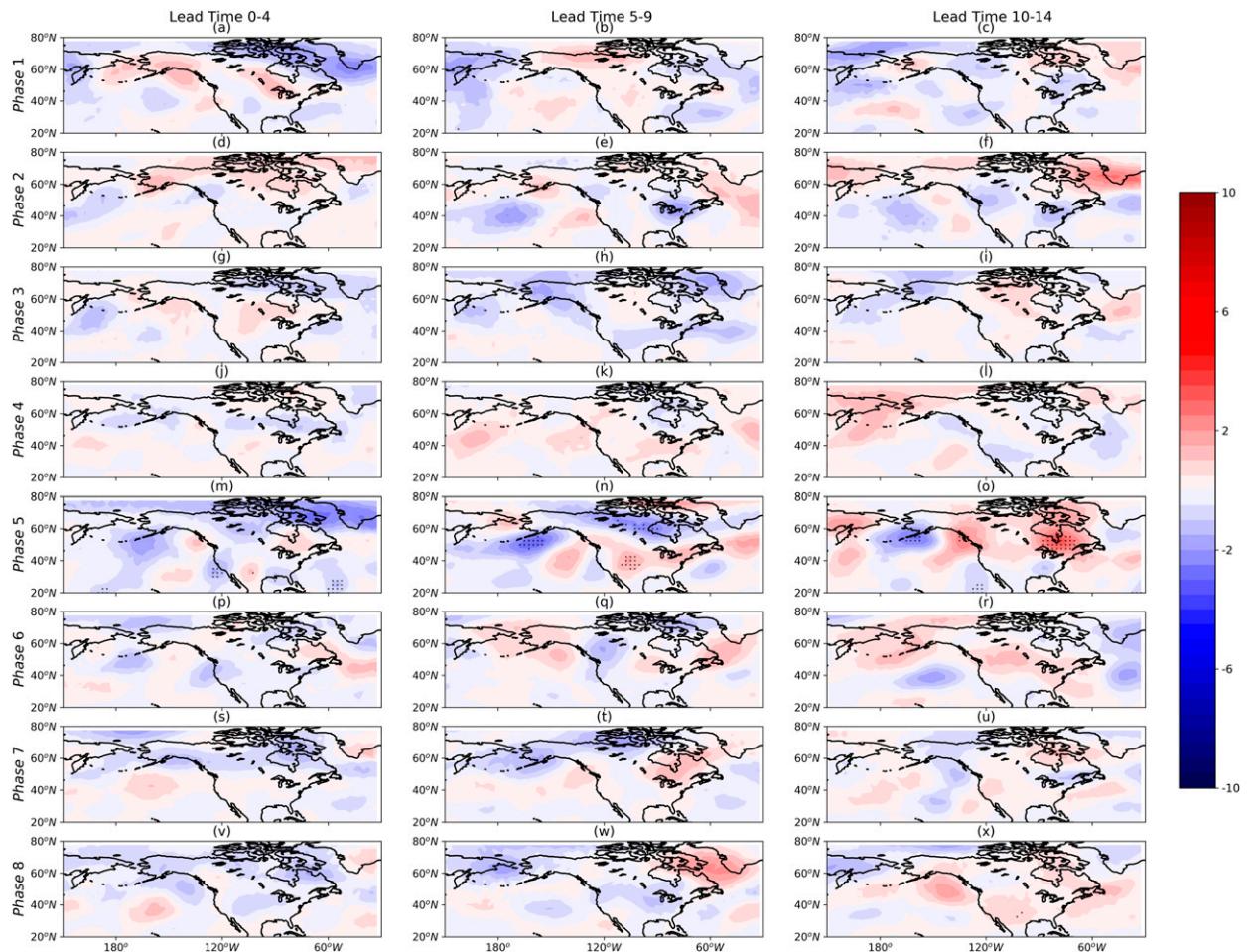


FIG. B3. As in Fig. B2, but for $k = -20$ minus $k = -15$.

REFERENCES

- Adames, Á. F., and D. Kim, 2016: The MJO as a dispersive, convectively coupled moisture wave: Theory and observations. *J. Atmos. Sci.*, **73**, 913–941, <https://doi.org/10.1175/JAS-D-15-0170.1>.
- Baggett, C. F., K. M. Nardi, S. J. Childs, S. N. Zito, E. A. Barnes, and E. D. Maloney, 2018: Skillful subseasonal forecasts of weekly tornado and hail activity using the Madden–Julian Oscillation. *J. Geophys. Res. Atmos.*, **123**, 12 661–12 675, <https://doi.org/10.1029/2018JD029059>.
- Cassou, C., 2008: Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. *Nature*, **455**, 523–527, <https://doi.org/10.1038/nature07286>.
- Ching, L., C.-H. Sui, and M.-J. Yang, 2010: An analysis of the multiscale nature of tropical cyclone activities in June 2004: Climate background. *J. Geophys. Res.*, **115**, D24108, <https://doi.org/10.1029/2010jd013803>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- De Souza, E. B., and T. Ambrizzi, 2006: Modulation of the intraseasonal rainfall over tropical Brazil by the Madden–Julian Oscillation. *Int. J. Climatol.*, **26**, 1759–1776, <https://doi.org/10.1002/joc.1331>.
- Fauchereau, N., B. Pohl, and A. Lorrey, 2016: Extratropical impacts of the Madden–Julian Oscillation over New Zealand from a weather regime perspective. *J. Climate*, **29**, 2161–2175, <https://doi.org/10.1175/JCLI-D-15-0152.1>.
- Hamill, T.-M., and G.-N. Kiladis, 2014: Skill of the MJO and Northern Hemisphere blocking in GEFS medium-range reforecasts. *Mon. Wea. Rev.*, **142**, 868–885, <https://doi.org/10.1175/MWR-D-13-00199.1>.
- Henderson, S. A., E. D. Maloney, and S. Son, 2017: Madden–Julian Oscillation Pacific teleconnections: The impact of the basic state and MJO representation in general circulation models. *J. Climate*, **30**, 4567–4587, <https://doi.org/10.1175/JCLI-D-16-0789.1>.
- Kiladis, G. N., J. Dias, K. H. Straub, M. C. Wheeler, S. N. Tulich, K. Kikuchi, K. M. Weickmann, and M. J. Ventrice, 2014: A comparison of OLR and circulation-based indices for tracking the MJO. *Mon. Wea. Rev.*, **142**, 1697–1715, <https://doi.org/10.1175/MWR-D-13-00301.1>.
- Lawrence, D. M., and P. J. Webster, 2002: The boreal summer intraseasonal oscillation: Relationship between northward and eastward movement of convection. *J. Atmos. Sci.*, **59**,

- 1593–1606, [https://doi.org/10.1175/1520-0469\(2002\)059<1593:TBSIOR>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<1593:TBSIOR>2.0.CO;2).
- Lin, H., 2018: Predicting the dominant patterns of subseasonal variability of wintertime surface air temperature in extratropical Northern Hemisphere. *Geophys. Res. Lett.*, **45**, 4381–4389, <https://doi.org/10.1029/2018GL077509>.
- Madden, R. A., and P.-R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708, [https://doi.org/10.1175/1520-0469\(1971\)028<0702:DOADOI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2).
- Maloney, E. D., and D. L. Hartmann, 2000: Modulation of hurricane activity in the Gulf of Mexico by the Madden–Julian Oscillation. *Science*, **287**, 2002–2004, <https://doi.org/10.1126/science.287.5460.2002>.
- Moore, A. M., and R. Kleeman, 1999: Stochastic forcing of ENSO by the intraseasonal oscillation. *J. Climate*, **12**, 1199–1220, [https://doi.org/10.1175/1520-0442\(1999\)012<1199:SFOEBT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<1199:SFOEBT>2.0.CO;2).
- Mori, M., and M. Watanabe, 2008: The growth and triggering mechanisms of the PNA: A MJO–PNA coherence. *J. Meteor. Soc. Japan*, **86**, 213–236, <https://doi.org/10.2151/jmsj.86.213>.
- Mundhenk, B. D., E. A. Barnes, and E. D. Maloney, 2016: All-season climatology and variability of atmospheric river frequencies over the North Pacific. *J. Climate*, **29**, 4885–4903, <https://doi.org/10.1175/JCLI-D-15-0655.1>.
- , —, —, and C. F. Baggett, 2018: Skillful empirical subseasonal prediction of landfalling atmospheric river activity using the Madden–Julian Oscillation and quasi-biennial oscillation. *npj Climate Atmos. Sci.*, **1**, 20177, <https://doi.org/10.1038/S41612-017-0008-2>.
- Seo, K.-H., and H.-J. Lee, 2017: Mechanisms for a PNA-like teleconnection pattern in response to the MJO. *J. Atmos. Sci.*, **74**, 1767–1781, <https://doi.org/10.1175/JAS-D-16-0343.1>.
- Slade, S. A., and E. D. Maloney, 2013: An intraseasonal prediction model of Atlantic and east Pacific tropical cyclone genesis. *Mon. Wea. Rev.*, **141**, 1925–1942, <https://doi.org/10.1175/MWR-D-12-00268.1>.
- Tseng, K.-C., E. A. Barnes, and E. D. Maloney, 2018: Prediction of the midlatitude response to strong Madden–Julian Oscillation events on S2S timescales. *Geophys. Res. Lett.*, **45**, 463–470, <https://doi.org/10.1002/2017GL075734>.
- , E. Maloney, and E. Barnes, 2019: The consistency of MJO teleconnection patterns: An explanation using linear Rossby wave theory. *J. Climate*, **32**, 531–548, <https://doi.org/10.1175/JCLI-D-18-0211.1>.
- Watanabe, M., and M. Kimoto, 2000: Atmosphere–ocean thermal coupling in the North Atlantic: A positive feedback. *Quart. J. Roy. Meteor. Soc.*, **126**, 3343–3369, <https://doi.org/10.1002/qj.49712657017>.
- Wheeler, M. C., H. H. Hendon, S. Cleland, H. Meinke, and A. Donald, 2009: Impacts of the Madden–Julian oscillation on Australian rainfall and circulation. *J. Climate*, **22**, 1482–1498, <https://doi.org/10.1175/2008JCLI2595.1>.
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. Elsevier, 676 pp.
- Yanai, M., S. Esbensen, and J.-H. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.*, **30**, 611–627, [https://doi.org/10.1175/1520-0469\(1973\)030<0611:DOBPOT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1973)030<0611:DOBPOT>2.0.CO;2).